



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

New insights into Parkinson's disease through statistical analysis of standard clinical scales quantifying symptom severity

Citation for published version:

Tsanas, A 2019, New insights into Parkinson's disease through statistical analysis of standard clinical scales quantifying symptom severity. in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Institute of Electrical and Electronics Engineers (IEEE), 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society , Berlin, Germany, 23/07/19. <https://doi.org/10.1109/EMBC.2019.8856559>

Digital Object Identifier (DOI):

[10.1109/EMBC.2019.8856559](https://doi.org/10.1109/EMBC.2019.8856559)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



New insights into Parkinson’s disease through statistical analysis of standard clinical scales quantifying symptom severity

Athanasios Tsanas

Abstract— Clinical research studies in Parkinson’s Disease (PD) focusing on symptom assessment often rely on thorough time-consuming physical examinations quantified on clinical scales such as the Unified Parkinson’s Disease Rating Scale (UPDRS). Although widely used in clinical research, realistic time constraints preclude its use in daily clinical practice. The Hoehn and Yahr (H&Y) staging is an alternative scale which is easier to administer and provides a succinct descriptor of overall PD severity. There is no universal agreement amongst neurologists on the specific PD symptoms they need to be assessing in order to prescribe treatments and optimize symptom management for their patients, and practically there are no clinical scales which are recorded in daily clinical practice. In this study, we systematically evaluate diverse symptoms (as expressed in 44 UPDRS items) and aim to provide a statistical association with UPDRS and H&Y using rank correlation and mutual information metrics. Moreover, we investigate the projection of a UPDRS item subset on a 2D plot to map onto H&Y. We report some statistically strong correlations of PD symptoms against UPDRS and H&Y ($|R| \geq 0.3$), and provide an intuitively appealing visualization mapping onto H&Y. These findings may be useful to neurologists as practical guidance in their daily clinical routine.

I. INTRODUCTION

Parkinson’s Disease (PD) is a chronic neurodegenerative disorder characterized by the manifestation of a range of motor symptoms and non-motor symptoms [1]. Disease onset is likely a combination of dopaminergic neuron reduction in the basal ganglia and degeneration of non-dopaminergic pathways [1]. Medication and surgical intervention can alleviate most of the symptoms and improve quality of life for most People With Parkinson’s (PWP) [2]. To optimize treatment frequent physical assessment by expert neurologists is required, however, in practice PWP are typically followed up by expert clinicians at intervals often spanning 6 or 12 months, which may underestimate symptom severity progression.

Standardized *clinical scales* have been proposed to quantify PD symptom severity. The Unified Parkinson’s Disease Rating Scale (UPDRS) [3] is the most widely used scale in clinical research studies [4], quantifying the constellation of multiple *motor*, *non-motor* PD symptoms, and complications of therapy. It consists of 44 *items*, where each item spans the range 0 (symptom-free) to 4 (severe disability). It is organized around four major *components*, each composed of a number of items: (1) Mentation, behavior and mood (4 items, items 1-4); (2) Activities of Daily Living (ADL, 13 items, items 5-17), assessing whether PWP can

complete daily tasks unassisted; (3) Motor (27 items, items 18-44), addressing muscular control; and (4) Complication of therapy (11 items, items 45-55), which expresses dyskinesia and disability problems associated with PD treatment. The third component is commonly referred to as *motor UPDRS*, and is highly correlated with the total UPDRS [5,6]. For untreated PWP the fourth component is not used. The total UPDRS is computed by summing up all 44 items (range: 0-176) and the motor UPDRS by summing up items 18-44 (range: 0-108).

Although UPDRS is the gold standard in PD research studies, in clinical practice it is rather seldomly used due to pragmatic time-constraints. Clinicians anecdotally use subsets of UPDRS, focusing on specific items that in their opinion are best suited to assess symptoms, and often only the motor UPDRS component is administered. Currently, there is no uniquely defined subset of items or agreement amongst expert neurologists on the most suitable time-efficient approach to assess PD symptom severity in daily clinical practice.

Alternative clinical scales have also been proposed, mainly to provide succinct, rudimentary descriptors of overall PD symptom severity. Hoehn and Yahr (H&Y) staging [7] is particularly practical and especially the *modified H&Y* is commonly used in many clinical centres [8]. The modified H&Y is an ordinal scale of eight possible *stages*: 0, 1, 1.5, 2, 2.5, 3, 4, 5, where increasing stage progressively quantifies more debilitating PD symptoms, providing a simple, general clinical impression of PD which requires a less thorough physical examination than UPDRS.

The motivation of this study is to elucidate the use of empirical practical tests used in clinical settings for PD assessment with the two most commonly used clinical PD research scales, UPDRS and H&Y staging. This study’s findings provide insight into (a) the statistical association of empirical practical tests which may be used as a proxy for assessing overall symptom severity (b) the statistical association of UPDRS and H&Y, and (c) the UPDRS items which may be most predictive of the clinical scales. We envisage these findings may be useful to PD neurologists facilitating their daily routine assessments in overall PD overview assessment using a subset of UPDRS items.

II. DATA

The database source of this study is the Parkinson’s Disease Data and Organizing Center (PD-DOC) [9] which was developed to facilitate the planning, study design, and

* The early part of this work was supported by the Wellcome Trust through a Centre Grant No. 098461/Z/12/Z, “The University of Oxford Sleep and Circadian Neuroscience Institute (SCNi)”.

A. Tsanas is with the Usher Institute of Population Health Sciences and Informatics, Medical School, University of Edinburgh, and with the Oxford Centre for Industrial and Applied Mathematics, Mathematical Institute, University of Oxford. (phone: +44 1316517887; email: atsanas@ed.ac.uk)

statistical analysis of PD-related data. The cohort is comprised of 566 PWP (1623 samples because some PWP had more than one evaluation), collected by several independent clinical centers, organized by the University of Rochester. Ethical approval was obtained by the IRBs of the participating centers. After removing samples containing missing UPDRS and H&Y entries, we processed 1594 samples from 562 PWP (363 males). The participants' age was (median \pm interquartile range): 63 \pm 14 years, with total UPDRS score 28 \pm 20, and H&Y stage 2 \pm 0. For further details see [9].

III. METHODS

A. Statistical analysis

We use rank correlation analysis to quantify the association strength of UPDRS items with the modified H&Y stages. We assess statistical significance (at $p = 0.01$ level) of the null hypothesis that UPDRS items have no correlation with total UPDRS or H&Y staging. The strength of this association is expressed by the magnitude of the correlation coefficient R . We followed the standard empirical rule for clinical applications to denote relationships as *statistically strong* when the magnitude of the correlation coefficient is equal or above 0.3 [10]. We also computed the *Mutual Information* (MI) [11] which quantifies the statistical dependence of H&Y on the UPDRS items (MI can account for general *nonlinear* relationships between variables). We report the *normalized* MI on the scale 0 to 1, with 0 indicating no statistical dependence between random variables, using our implementation [6].

B. Data projection and visualization in 2D embedded space

Clinical questionnaires and clinical scales often exhibit some underlying hidden structure: the rationale is that they capture some intrinsic characteristics which are not directly observed or quantified. This might be usefully presented in the embedded space, where visualization might lead to new insights. A widely used approach to project data in 2D is t-distributed Stochastic Neighbor Embedding (t-SNE) [12]. Here, we used t-SNE to visualize how UPDRS items can be projected in the 2D space mapped onto H&Y.

IV. RESULTS

Figures (1) and (2) present scatter plots to visualize the statistical relationships of UPDRS items with UPDRS and H&Y. The statistical associations between UPDRS items and total UPDRS are summarized in Table I, and between UPDRS items and H&Y are summarized in Table II. In all cases, the relationships are statistically significant ($p < 0.001$). These findings suggest that bradykinesia (item 44) followed by ADL items (dressing and eating food) are the most statistically strongly associated symptoms towards estimating total UPDRS. Similarly, item 43 (postural stability) is, statistically, the most important UPDRS item in determining H&Y, followed by items 42 (gait), 40 (arise from chair), and 44 (bradykinesia), which is in broad agreement with clinical expectations. Interestingly, the postural stability item exhibits markedly larger association strength compared to the other UPDRS items in terms of H&Y association. Finally, Fig. 3 presents the projection of the 10 UPDRS items from Table II onto H&Y: this plot visually indicates that a subset of UPDRS items may provide an overall differentiation of the different H&Y stages, particularly for H&Y stages 4 and 5 (marking substantial symptom severity).

TABLE I. ASSOCIATION STRENGTH (NORMALIZED MUTUAL INFORMATION & RANK CORRELATIONS) OF ITEMS WITH TOTAL UPDRS

| All data ($n = 1594$ samples) | | |
|--------------------------------|-------|----------------------|
| UPDRS item | MI | Spearman correlation |
| (44) Bradykinesia | 0.089 | 0.621 |
| (10) Dressing | 0.077 | 0.60 |
| (9) Eating food | 0.074 | 0.581 |
| (19) Face expression | 0.073 | 0.568 |
| (35) Hand move LH | 0.071 | 0.583 |
| (11) Hygiene | 0.067 | 0.539 |
| (18) Speech expression | 0.067 | 0.558 |
| (27) Rigid neck | 0.066 | 0.533 |
| (34) Hand move RH | 0.066 | 0.559 |
| (36) RAM RH | 0.065 | 0.556 |

All reported UPDRS items were statistically significantly correlated with the total UPDRS ($p < 0.001$). For clarity, only the 10 UPDRS items most strongly associated with UPDRS are presented here. We report the normalized mutual information (MI, range 0-1). LH stands for 'left hand', 'RH' for right hand, and RAM for 'rapid alternating movement'. The numbers in parenthesis refer to the UPDRS item order.

TABLE II. ASSOCIATION STRENGTH (NORMALIZED MUTUAL INFORMATION & RANK CORRELATIONS) OF UPDRS ITEMS WITH H&Y

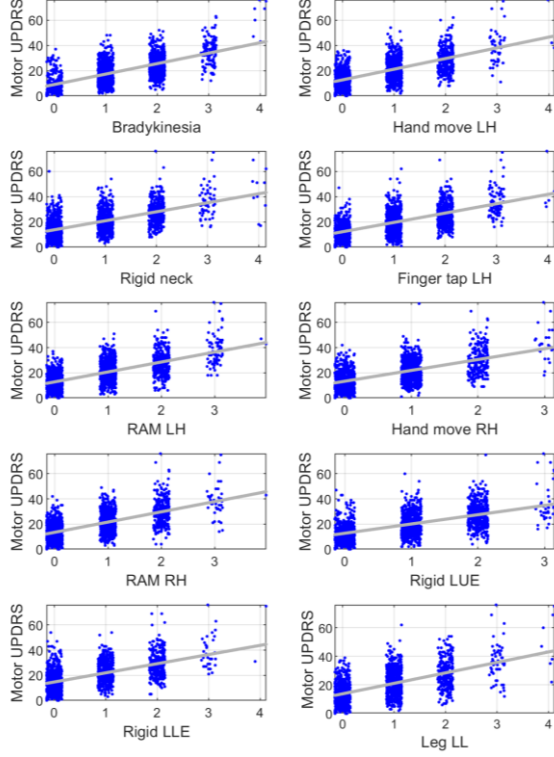
| All data ($n = 1594$ samples) | | |
|--------------------------------|-------|----------------------|
| UPDRS item | MI | Spearman correlation |
| (43) Postural stability | 0.334 | 0.614 |
| (42) Gait | 0.157 | 0.403 |
| (40) Arise from chair | 0.135 | 0.379 |
| (44) Bradykinesia | 0.132 | 0.413 |
| (15) Walking | 0.120 | 0.338 |
| (41) Posture | 0.104 | 0.388 |
| (12) Turn bed | 0.086 | 0.337 |
| (10) Dressing | 0.084 | 0.334 |
| (19) Face expression | 0.084 | 0.342 |
| (11) Hygiene | 0.083 | 0.290 |

All reported UPDRS items were statistically significantly correlated with the modified H&Y stages ($p < 0.001$). For clarity, only the 10 UPDRS items most strongly associated with H&Y are presented here. We report the normalized mutual information (MI, range 0-1). The numbers in parenthesis refer to the UPDRS item order.

V. DISCUSSION

This study provided a systematic investigation into widely used empirical practical tests that clinicians use in their routine PD assessment to estimate their statistical association with total UPDRS and the modified H&Y staging. We used the standard quantification through UPDRS items as a convenient means of expressing symptom severity. In a previous study we optimized a parametric functional mapping of motor UPDRS onto H&Y [13], when all motor UPDRS items are available. Here, we tackled a more generic question attempting to provide overall assessment of PD symptom severity (quantified using total UPDRS and H&Y) without the need of recording the full span of the UPDRS items. This is a particularly pertinent question, given that typical physical examinations of PD patients anecdotally last about 20 minutes whereas the total UPDRS assessment administration requires longer (depending on the expert rater's experience).

Scatter plot (with jitter): UPDRS items vs motor UPDRS



Scatter plot (with jitter): UPDRS items vs total UPDRS

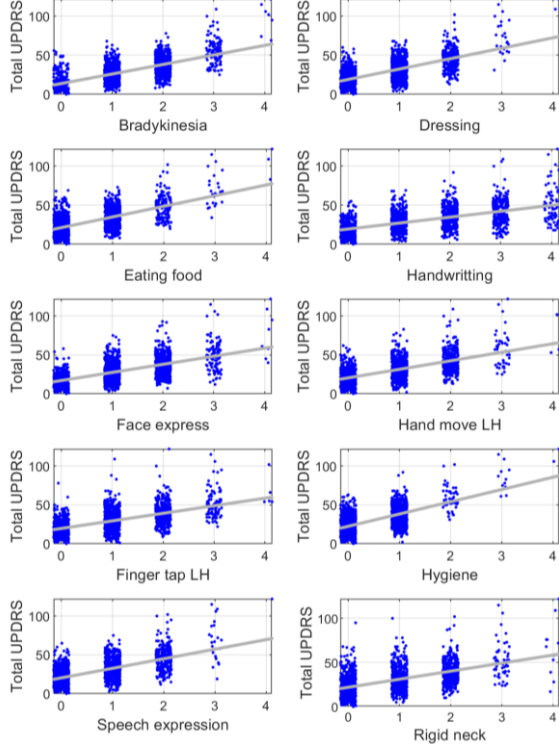


Fig. 1. Scatter plots of the most strongly associated UPDRS items with motor UPDRS (top), and total UPDRS (bottom). The robust regression line indicates the trend in the data [6], and we have added 15% jitter on the UPDRS items to visualize better the number of overlapping points. LH stands for ‘left hand’, ‘RH’ for ‘right hand’, LUE for ‘left upper extremity’, LLE for left lower extremity’, and LL for ‘left leg’.

Scatter plot (with jitter): UPDRS items vs modified H&Y

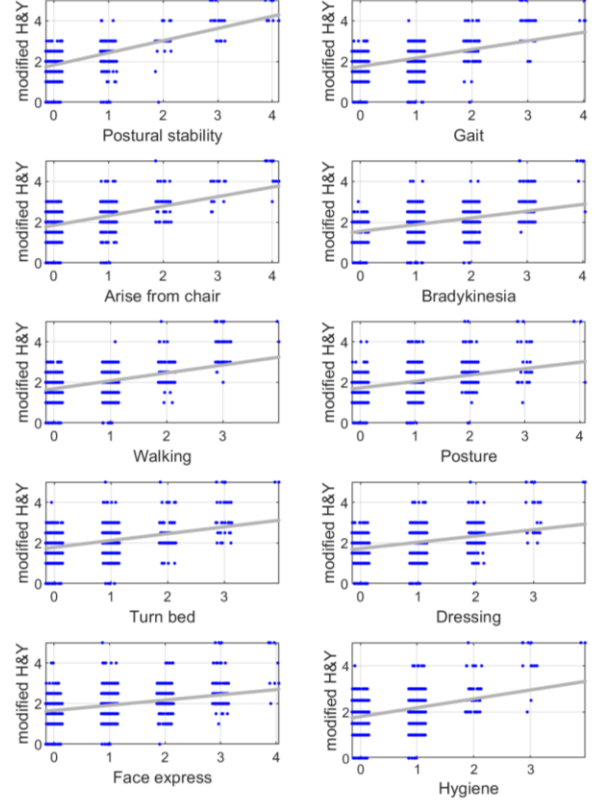


Fig.2 Scatter plots of the most strongly associated UPDRS items with modified H&Y. The robust regression line indicates the trend in the data [6], and we have added 15% jitter on the UPDRS items to visualize better the number of overlapping points.

The findings in Table I provide very clear guidance on the UPDRS item which are univariately most strongly associated with the total UPDRS. We computed very strong statistical relationships ($|R| \geq 0.3$) across multiple items. We remark that some are part of motor UPDRS (bradykinesia, face expression and speech expression), whilst there are some items from the activities of daily living component (dressing, eating food, and hygiene). These insights serve well to highlight the kind of symptoms clinicians may want to place more emphasis on when doing their routine assessments.

The results of the statistical analysis in Table II suggest that UPDRS items 40 (posture), 42 (gait), 43 (postural stability) and 44 (bradykinesia) are most strongly correlated with H&Y stages. This finding is broad agreement with the revised optimized parametric formula we had proposed [13] which built on expert neurologists’ clinical intuition [14]. We remark that the UPDRS items most strongly associated with H&Y are indeed those from the motor component (items 18-44). Interestingly, item 15 (walking, part of the activities of daily living component) is one of the very few items that does not form part of the motor UPDRS component and yet is quite strongly associated with H&Y. However, this item can also be interpreted as referring to a motor symptom nonetheless. Overall, these results support the commonly-held notion that H&Y is predominantly sensitive to motor symptoms.

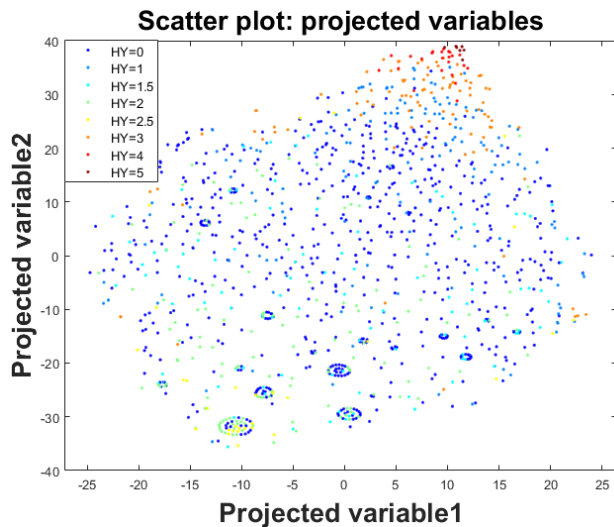


Fig.3. Projection of the 10 UPDRS items from Table II using t-SNE to visually assess the relationship with H&Y.

Fig. 3 projects the multi-dimensional UPDRS feature space in a visually appealing 2D plot. In a similar study, we previously investigated the constellation of samples for different symptom severity categories extracted through clustering [15]; here we directly visually assess the overall differentiation of the H&Y stages in the projected 2D space.

Stebbins and Goetz [16] used factor analysis to determine the motor UPDRS structure. They reported motor UPDRS can be assessed on six distinct factors: speech, facial expression, balance and gait (factor I), rest tremor (factor II), rigidity (factor IV), right and left bradykinesia (factors III and V), and postural tremor (factor VI). We had previously reported generally good agreement on a different dataset [6]. However, standard factor analysis does not offer a unique representation of the data and makes some strong underlying assumptions [17]. Future work could investigate the UPDRS latent variable structure using sparse principal component analysis which has provided interesting insights in a related study [18]. Determining the latent variables of UPDRS or ideally a sparse setting with few UPDRS items could subsequently facilitate analysis e.g. to assess PD symptom severity [5] and for early PD biomarkers [19], thus helping scale up research work when the entire UPDRS is not fully administered in standard clinical practice. This is very well aligned with our recent work aiming to develop tools that can be deployed at large scale in community studies [20].

ACKNOWLEDGMENT

The data source for this work was the Parkinson's Disease Data and Organizing Center (PD-DOC; U01NS050095, PI: Roger Kurlan), currently managed by the National Institute of Neurological Disorders and Stroke (NINDS). Data has been contributed to the PD-DOC by the NINDS Udall Centers of Excellence for Parkinson's Disease Research, the NINDS Neuroprotective Exploratory Trials in Parkinson's Disease Program and the Parkinson Study Group. We used data from two studies available in the original PD-DOC repository: (1) the study from three Udall Centers (PIs: (i) T.M. Dawson, (ii) M.K. Lee, (iii) V.L. Dawson), and (2) the PostCEPT study (PIs: B. Ravina and A.E. Lang).

REFERENCES

- [1] C.W. Olanow, M.B. Stern, K. Sethi, "The scientific and clinical basis for the treatment of Parkinson disease", *Neurology*, Vol. 72, pp. s1-s136, 2009
- [2] Singh, N, Pillay, V, Choonara E: "Advances in the treatment of Parkinson's disease", *Progress in Neurobiology* 2007; 81:29-44
- [3] Fahn S, Elton RL, Members of the UPDRS Development Committee. Unified Parkinson's disease rating scale. In: Fahn S, Marsden CD, Goldstein M, et al, eds. *Recent developments in Parkinson's disease*. Vol II. New Jersey: Macmillan Healthcare Information, 1987:153-63.
- [4] Ramaker, C, Marinus, J, Stiggelbout AM, van Hilten BJ. Systematic evaluation of rating scales for impairment and disability in Parkinson's disease, *Movement Disorders* 2002; 17:867-876
- [5] A. Tsanas, M.A. Little, P.E. McSharry, L.O. Ramig: "Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity", *Journal Royal Society Interface* 2011; 8:842-855
- [6] A. Tsanas, "Accurate telemonitoring of Parkinson's disease symptom severity using nonlinear speech signal processing and statistical machine learning", PhD thesis, Oxford Centre for Industrial and Applied Mathematics, University of Oxford, 2012
- [7] Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality, *Neurology* 1967; 17(5):427-42
- [8] Goetz CG, Poewe W, Rascol O, Sampaio C, Stebbins GT, Counsell C, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations, *Movement Disorders* 2004; 19(9):1020-8
- [9] R. Kurlan R, D. Murphy: "Parkinson's disease data and organizing center", *Movement Disorders* 2007; 22:904-5
- [10] A. Tsanas, M.A. Little, P.E. McSharry: "A methodology for the analysis of medical data", in *Handbook of Systems and Complexity in Health*, Eds. J.P. Sturmberg, and C.M. Martin, Springer, pp. 113-125 (chapter 7), 2013
- [11] T.M. Cover, J.A. Thomas, *Elements of information theory*, 2nd ed., Wiley, ISBN: 978-0-471-24195-9, 2006
- [12] L.J.P. van der Maaten, G.E. Hinton, "Visualizing data using t-SNE", *Journal of Machine Learning Research* 2008; 9:2579-2605
- [13] A. Tsanas, M.A. Little, P.E. McSharry, B.K. Scanlon, S. Papapetropoulos: "Statistical analysis and mapping of the Unified Parkinson's Disease Rating Scale to Hoehn and Yahr staging", *Parkinsonism and Related Disorders* 2012; 18 (5):697-699
- [14] B.K. Scanlon, H.L. Katzen, B.E. Levin, C. Singer, S. Papapetropoulos, "A revised formula for the conversion of UPDRS-III scores to Hoehn and Yahr stage", *Parkinsonism and Related Disorders* 2010; 16: 151-2
- [15] B. Sheaves, K. Porcheret, A. Tsanas, C. Espie, R. Foster, D. Freeman, P.J. Harrison, K. Wulff, G.M. Goodwin: "Insomnia, nightmares, and chronotype as markers of risk for severe mental illness: results from a student population", *Sleep* 2016; 39(1):173-181
- [16] G. Stebbins, C. Goetz, "Factor Structure of the Unified Parkinson's Disease Rating Scale: Motor Examination Section", *Movement Disorders*, 1998; 13:633-636
- [17] T. Hastie, R. Tibshirani, J. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, 2nd edition, Springer, New York, USA, 2009
- [18] A. Tsanas, K.E.A. Saunders, A.C. Bilderbeck, N. Palmius, G.M. Goodwin, M. De Vos: "Clinical insight into latent variables of psychiatric questionnaires for mood symptom self-assessment", *JMIR Mental Health*, 2017; 4(2):e15
- [19] S. Arora, N.P. Visnaji, T.A. Mestre, A. Tsanas, A. Al Dakheel, B.S. Connolly, C. Gasca-Salas, D.S. Kern, J. Jain, E.J. Slow, A. Faust-Socher, A.E. Lang, M.A. Little, C. Marras: "Investigating voice as a biomarker for leucine-rich repeat kinase 2-associated Parkinson's disease: a pilot study", *Journal of Parkinson's Disease* 2018; 8:503-510
- [20] S. Arora, L. Baghai-Ravary, A. Tsanas: "Developing a large scale population screening tool for the assessment of Parkinson's disease using telephone-quality speech", *Journal of Acoustical Society of America*, *in press*